

ISSUES , IDEAS
AND
INFORMATION
FOR PSYCHOLOGY
STUDENTS

NO.5 - RESEARCH
METHODS

Kevin Brewer

ISBN: 978-1-904542-43-8

This document is produced under two principles:

1. All work is sourced to the original authors. The images are all available in the public domain (most from http://commons.wikimedia.org/wiki/Main_Page). You are free to use this document, but, please, quote the source (Kevin Brewer 2009) and do not claim it as you own work.

This work is licensed under the Creative Commons Attribution (by) 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 171 2nd Street, Suite 300, San Francisco, California, 94105, USA.

2. Details of the author are included so that the level of expertise of the writer can be assessed. This compares to documents which are not named and it is not possible to tell if the writer has any knowledge about their subject.

Kevin Brewer BSocSc, MSc

An independent academic psychologist, based in England, who has written extensively on different areas of psychology with an emphasis on the critical stance towards traditional ideas.

Orsett Psychological Services,
PO Box 179,
Grays,
Essex
RM16 3EW
UK

orsettpsychologicalservices@phonecoop.coop

(<http://kmbpsychology.jottit.com>) or
(<http://psyman.weebly.com>)

CONTENTS

	Page Number
1. The Single Participant Experiment and Blindsight	4
2. Measuring Non-Conscious Perception	8
3. Recruiting Volunteer Participants	11
4. Descriptive Statistics: Averages and Measures of Central Tendency	17
5. Cross-sectional Longitudinal Studies: The Best of Both Worlds	27
6. Studying Behaviour with an Internet-Based Questionnaire: An Example from Research on Sleep Disorders	32
7. Studying Specific Groups: The Nun Study	35
8. Psychology Debate: This House Believes that the Self is Best Studied with Quantitative Methods	38
9. Two Classic 1960s Social Psychology Studies on the Influence of Other People: Laboratory versus Naturalistic Study	40
10. Carving Up Continuous Data	50

1. THE SINGLE PARTICIPANT EXPERIMENT AND BLINDSIGHT

- 1.1. Introduction
- 1.2. GY and blindsight
- 1.3. References

1.1. INTRODUCTION

The single participant experiment is not a case study. It is an experiment, with all the rigour and control, that uses a single participant. The person in question is typically unusual and thus the benefits of studying them (table 1.1).

Though it appears to be similar to the repeated measures design, it is technically classed as an unrelated design, particular for the statistical test used (Coolican 1998). The only requirement is that the scores are obtained independent of each other.

ADVANTAGES

1. Useful when few or "special" participants available.
2. Useful when a lot of time and/or cost required to prepare and/or train participants. GY is used to the techniques of the experiments used to test his blindsight.
3. Maintain rigour and control of the experimental method.
4. Better than a case study when full details of the participant's life is not required.

DISADVANTAGES

1. The results cannot be generalised to a wider population.
2. It is dependent on the participant because they drop-out, the experiment stops.
3. The participant knows they are being studied, thus risk of demand characteristics or evaluation apprehension.
4. The participant will be extra sensitive to the results as they are being studied alone.

Table 1.1 - Advantages and disadvantages of the single participant experiment.

There are a number of different designs for the single participant experiment (Wilson 1995):

- AB design - The participant does two conditions, one experimental and one control, and the scores are compared;

CONTROL → EXPERIMENTAL

- ABAB design - As above but the participant repeats the two conditions. This covers any risk that the scores in one condition were a fluke and gives two sets of scores for the control and the experimental condition to average;

CONTROL → EXPERIMENTAL → CONTROL → EXPERIMENTAL

- ABC design - The participant undertakes two different experimental conditions as well as the control condition.

CONTROL → EXPERIMENTAL 1 → EXPERIMENTAL 2

1.2. GY AND BLINDSIGHT

One individual who has been studied in detail in such experiments is GY. He is a middle-aged man, who due to a childhood brain injury, is blind in his right visual field. His eyes are unaffected, but the left visual cortex is damaged (information from the right eye goes to the left hemisphere and vice versa). However, he is still able to "see" things in the right visual field though he has no conscious awareness of them. This is known as blindsight (Weiskrantz 1996).

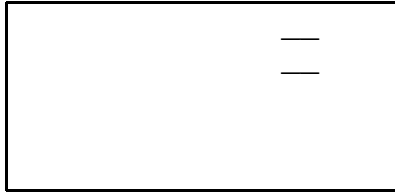
Kentridge et al (1999) were interested to know if GY could selectively attend to objects in his blind field. So they tested him in a single participant experiment involving direct and indirect cueing (techniques developed by Posner 1980).

Central cueing involves participants detecting a target on a computer screen as quickly as possible. Before the target is presented, an arrow appears briefly pointing to an area of the screen. If this arrow is cueing the correct area for the target, reaction time to detect the target is quicker than if there is no arrow or when the arrow cues the wrong area. The arrow appears for 0.6 seconds, then an 0.5 second interval with an auditory tone before the target is presented for 0.5 seconds (Kentridge et al 1999).

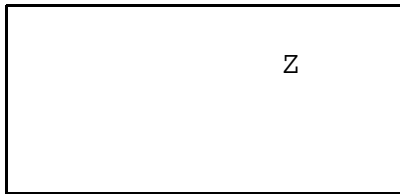
Direct peripheral cueing uses a set of parallel lines instead of an arrow. Indirect peripheral cueing

involves the parallel lines appearing where the target is unlikely to be (figure 1.1). The purpose of the cue is to focus attention to part of the screen.

Screen (A)



Screen (B)



Screen (C)



Direct peripheral cueing = cue (A) followed by (B)
Indirect peripheral cueing = cue (A) followed by (C)

Figure 1.1 - Direct and indirect peripheral cueing.

In each condition, the target was in the area cued 68.75% of trials and not in 31.25% of trials.

GY was tested in ten blocks of 128 trials for central cueing, and twenty blocks of 128 trials for indirect peripheral cueing.

The results showed that GY was significantly quicker to detect targets when the cue was correct in the central cueing condition and more accurate (69% correct) despite consciously not seeing anything. he said, for example, "I would be none the wiser if you were not putting any cues up just to confuse me", and "I just listen for the beep and press a button" (p1808).

He was significantly slower when the cue was misleading. The same results occurred for direct peripheral cueing. With the indirect peripheral cueing condition, he was better at detecting the target with the misleading cues. All the results suggested that he was focusing his attention with the cue. This would seem to

be attention without awareness.

1.3. REFERENCES

Coolican, H (1998) Teaching psychology methods and statistics at A(ny) level. In Jarvis, M (ed) 1998 ATP Conference Leicester: Association for the Teaching of Psychology

Kentridge, R.W et al (1999) Attention without awareness in blindsight Proceedings of the Royal Society London 266, 1805-1811

Posner, M.I (1980) Orienting of attention Quarterly Journal of Experimental Psychology 32, 3-25

Weiskrantz, L (1996) Blindsight revisited Current Opinion in Neurobiology 6, 215-220

Wilson, S.L (1995) Single case designs. In Breakwell, G.M et al (eds) Research Methods in Psychology London: Sage

2. MEASURING NON-CONSCIOUS PERCEPTION

- 2.1. Introduction
- 2.2. Tamietto and de Gelder (2008)
- 2.3. References

2.1. INTRODUCTION

Non-conscious processing has been tested using a technique called backward masking. Two stimuli are presented, one after the other, but the first was too quick to be consciously seen. The first stimulus (target) will still affect the participant even if they can only consciously see the second stimulus (mask). For example, a picture of an angry face followed by a neutral face. The stimulus onset asynchrony (SOA)(time between the onset of the target and the mask) is typically less than 40 milliseconds (Tamietto and de Gelder 2008). The target, though not consciously seen, does have an effect as measured by, for example, skin conductance response (speed of minute electrical crossing the surface of the skin used as measure of arousal).

2.2. TAMIETTO AND DE GELDER (2008)

Tamietto and de Gelder (2008) used backward masking with pictures of facial expressions. Twenty-five volunteers (14 women, 11 men) from Torino, Italy were recruited. The target face was presented for 20 milliseconds and the mask face (which was scrambled) for 180 milliseconds. The twelve faces were either showing a happy or a fearful expression depending on the condition (table 2.1). The researchers also varied the process for each eye.

	TARGET	MASK
Congruent A	Happy	Happy
Congruent B	Fearful	Fearful
Incongruent A	Happy	Fearful
Incongruent B	Fearful	Happy

Table 2.1 - Combination of faces in different conditions.

The task was to press a computer key as quickly as possible if the face (mask) showed a particular emotion. This is known as the Go/No Go technique (table 2.2).

	QUESTION	FACE SHOWN
"Go" condition	Is face happy?	Happy
"No Go" condition	Is face happy?	Fearful

Table 2.2 - Go/No Go technique.

Reaction time for the congruent conditions was faster (460ms on average) than for the incongruent conditions (480ms on average). This difference in reaction time, though small, is taken as evidence that the target interferes with the mask (ie: the target was perceived even if only at a non-conscious level). Furthermore, a non-conscious fearful face significantly slowed the reaction time for recognition of the happy face (incongruent B in table 2.1), but not the other way around (incongruent A).

This experiment has a number of strengths and weaknesses (table 2.3).

STRENGTHS

1. Allows for control and rigour that is typical of the experimental method. The standardised procedure included photographs of the same size, equal numbers of male and female faces, and the same time for the presentation of each one.
2. Reduces risk of demand characteristics with non-conscious stimuli. Participants naive to the backward masking procedure.
3. Reaction time can be measured objectively and accurately.
4. Computers allow automation of the experiment, which removes human influences from the procedure. For example, the stimuli were presented in the centre of the screen against a dark background on a 21-inch Sony CRT monitor. The monitor was connected to an IBM-compatible Pentium PC. Eye movements and eye tracking were also measured.
5. Repeated measure design was used. This reduces the influence of participant variables as the participant perform all conditions.
6. Full written informed consent was obtained from the participants and the research was approved beforehand by the Ethical Committee of the Department of Psychology, University of Torino, Italy.
7. The participants were checked beforehand for normal visual acuity and for no history of neurological or psychiatric illness.
8. A full pilot study with twenty similar participants was performed to choose the SOA from 10, 20, 30, 40, and 60ms.

WEAKNESSES

1. It was an artificial experiment (ie: not possible to test in real-life situation).

2. It was a complex experiment requiring computer equipment and laboratory facilities.
3. Participants were tested for many trials and may become tired or bored (and suffer from order effects). Order effects are where the performance on earlier trials influence performance on later trials. Four blocks of trials were run. Each block comprised 256 "Go" trials (positive recognition of facial expression) and 128 "No Go" trials (not facial expression shown).
4. Assumptions had to be made about the non-conscious stimuli as the participants reported not seeing it. This is a problem with indirect measures of behaviour.
5. The decision about the length of the SOA and how the faces scrambled were subjective.
6. Only a small number of participants were used, and the age ranged between 20-31 years only with slightly more women than men.
7. Individuals who volunteer for research may not be typical of the general population (Brewer 2005), which limits the generalisability of the findings.
8. There were design problems. For example, around 2% of trials were removed from analysis because of anticipation by the participants, delays, or participants reporting seeing the target face. The errors were around 17% (ie: misses in "Go" condition and false positives in "No Go" trials).

Table 2.3 - Strengths and weaknesses of the backward masking design and experiment of Tamietto and de Gelder (2008).

2.3. REFERENCES

Brewer, K (2005) The use of volunteers in psychological research
Psychology Information for Students 3, 8-11

Tamietto, M & de Gelder, B (2008) Affective blindsight in the intact brain: Neural interhemispheric summation for unseen fearful expressions
Neuropsychologia 46, 820-828

3. RECRUITING VOLUNTEER PARTICIPANTS

- 3.1. Introduction
- 3.2. Newspaper
- 3.3. Telephone/direct approach
- 3.4. Combination of techniques
- 3.5. References

3.1. INTRODUCTION

Using volunteers as participants in research is felt to be much better in terms of ethics (ie: right to non-participation) than the forced participation of, say, students as is/was common in many universities (Foot and Sanford 2004). Volunteers have advantages and disadvantages as participants (table 3.1).

ADVANTAGES

1. Allows right of non-participation which is more acceptable than forced participation.
2. In some cases only volunteers can be used; eg: studies of particular mental disorders, or "hidden"/deviant behaviour.
3. Student volunteers are a quick and convenient way to find participants for university researchers.

DISADVANTAGES

1. Individuals who volunteer are not necessarily typical of the general population; eg: more extravert; higher need for social approval than the norm (Brewer 2005).
2. It is a self-selecting sample. This produces samples than are unrepresentative of the population, and thus the results should not be generalised.
3. The motivations of the volunteers: "Volunteers go to the experimental setting because they want to; they have a vested interest in the outcome of the study, they see the research as important, or they are just curious. Any one of these concerns can prejudice the outcome of an experiment" (Nation 1997 p108).

Table 3.1 - Main advantages and disadvantages of using volunteers in psychological research.

Finding volunteers requires advertising the research project and asking for responses. Such advertising is done in a number of ways:

- Posters around a particular area like around a campus
- Newspaper

- Internet
- Word of mouth
- Telephone/direct approach

3.2. NEWSPAPER

Gunnarsson (2009) wanted healthy older adults for interviews about their everyday lives. Advertisements were placed in a regional newspaper in Sweden inviting those aged 75 years and older living in their own homes to apply. There were 25 replies of which five withdrew, leaving a study sample of twelve women and eight men. Table 3.2 summarises the characteristics of the sample.

	WOMEN	MEN
Number	12	8
Age range	75-90	75-89
Marital status	7 widowed 4 married	4 widowed 4 married
Former occupation	7 non-manual	6 non-manual

Table 3.2 - Characteristics of volunteer sample in Gunnarsson (2009).

In terms of reasons for volunteering:

They said that they took part in the study mainly because they thought it important to challenge the general impression of older people as frail and dependent, which they thought was dominant in the mass media and society. They simply wanted to provide evidence of lives that did not consist only of misery, and in this they thought they had something to tell" (p36).

Gunnarsson's research used qualitative indepth interviews, so the size and make-up of the sample was not as important as a quantitative study. The representativeness of the sample allows generalisation of the findings to the whole population, which is more common with quantitative studies.

In his classic study of obedience, Milgram (1963) recruited forty male volunteers from the New Haven area in the USA by newspaper and direct mail advertising. He asked for participants in a study on memory and learning at Yale University (figure 3.1).

Public Announcement

WE WILL PAY YOU \$4.00 FOR ONE HOUR OF YOUR TIME

Persons Needed for a Study of Memory

*We will pay five hundred New Haven men to help us complete a scientific study of memory and learning. The study is being done at Yale University.

*Each person who participates will be paid \$4.00 (plus 50c carfare) for approximately 1 hour's time. We need you for only one hour: there are no further obligations. You may choose the time you would like to come (evenings, weekdays, or weekends).

*No special training, education, or experience is needed. We want:

Factory workers

Businessmen

Construction workers

City employees

Clerks

Salespeople

Laborers

Professional people

White-collar workers

Barbers

Telephone workers

Others

All persons must be between the ages of 20 and 50. High school and college students cannot be used.

*If you meet these qualifications, fill out the coupon below and mail it now to Professor Stanley Milgram, Department of Psychology, Yale University, New Haven. You will be notified later of the specific time and place of the study. We reserve the right to decline any application.

*You will be paid \$4.00 (plus 50c carfare) as soon as you arrive at the laboratory.

TO:

PROF. STANLEY MILGRAM, DEPARTMENT OF PSYCHOLOGY,
YALE UNIVERSITY, NEW HAVEN, CONN. I want to take part in
this study of memory and learning. I am between the ages of 20 and
50. I will be paid \$4.00 (plus 50c carfare) if I participate.

NAME (Please Print)

ADDRESS

TELEPHONE NO. Best time to call you

(Source: Olivier Hammam; copied from original)

Figure 3.1 - Type of advertisement used by Milgram.

The volunteers in the original obedience experiment

were aged between 20 and 50: eight 20-29 years, 16 30-39 years, and 16 40-50 years old. In terms of occupation, nine were "professionals", sixteen "sales, business and white collar", and 15 were "workers skilled and unskilled" according to Milgram.

3.3. TELEPHONE/DIRECT APPROACH

Rather than waiting for volunteers to come forward, individuals can be directly contacted, usually by telephone, and asked to participate. There are advantages and disadvantages of this technique for recruiting participants (table 3.3).

ADVANTAGES	DISADVANTAGES
<p>1. Researchers do not have to wait for volunteers to come forward (ie: quicker way to gain sample).</p> <p>2. Allows the sample to be more diverse than other volunteer samples.</p> <p>3. Participants agree on the spot and the research takes place soon afterwards (or immediately if telephone interview). There is less chance of participants changing their minds and not</p>	<p>1. Individuals may feel obliged to participate because they cannot say no. Are these individuals really volunteering or they "arm-twisted volunteers"?</p> <p>2. "Arm-twisted volunteers" may not participate as fully/honestly as "full volunteers".</p> <p>3. It can involve a lot of effort (in terms of number of telephone calls made) for little reward (small sample).</p>

Table 3.3 - Advantages and disadvantages of a direct approach to recruiting participants.

Cheurprakobit (2000) recruited participants for his study about attitudes towards the police by randomly dialling telephones in the Midland and Odessa areas of Texas. Random numbers were generated which matched real telephone numbers. These numbers were then dialled by interviewers between 9am to 7pm on Mondays to Thursdays, and 10am to 4pm on Saturdays between May 19 and June 30, 1997. The random number technique produces a sampling pool quickly, and includes telephone numbers that may not be listed if telephone directories were used.

There were 2207 random calls made in Midland, which after exclusions became 115 individuals who had contacted the police in the past 24 months and were willing to be interviewed. In Odessa, 136 individuals were recruited from 2221 random calls (table 3.4). Overall, there was a lot of random dials to find an appropriate sample (5.7% of all dials).

	MIDLAND	ODESSA
Total ransom calls	2207	2221
Disconnected numbers	806	1062
Business numbers	297	197
Residential numbers:	<u>1104</u>	<u>962</u>
No answer/answering machine	569	448
No appropriate person/unable to communicate with interviewer	43	36
Refused to participate	205	184
Willing to participate but no police contact	172	158
Contacted police	115	136
% all dials	5.2	6.1

Table 3.4 - Breakdown of exclusions in Midland and Odessa.

The interviewers were instructed to randomise the order of asking for the youngest/oldest male/female over eighteen years in the household. The interviewers had four hours training, and had to be bilingual (English/Spanish) as there was a large Hispanic population in the area.

The results of the survey found ethnic differences in the attitudes towards the police, and a positive or negative attitude depended upon who initiated the contact (eg: citizen reporting crime or police arresting individual).

3.4. COMBINATION OF TECHNIQUES

Recruiting "hidden" populations needs multiple methods of advertising to find such volunteers. Miller et al (2007) was interested in lap dancer and their tip earnings over the menstrual cycle. He recruited eighteen women from gentlemen's clubs in New Mexico, USA by "indirect emails (forwarded through local industry contacts), newspaper advertisements, and flyers posted near clubs" (p377). Their mean age was 26.9 years, and the average length of lap dancing was 6.4 years.

One way to recruit individuals from a "hidden" or minority population is by snowball sampling. This begins with key people who introduce the researcher to others and the sample is built up this way. There are advantages

and disadvantages to snowball sampling (table 3.5).

ADVANTAGES	DISADVANTAGES
<ol style="list-style-type: none">1. It may be the only way to find participants when studying socially unacceptable or illegal behaviour.2. Good method when lists of potential participants are not available or contain omissions.3. Because the researcher is introduced by members of the population, others may be more willing to participate in the research than a direct approach	<ol style="list-style-type: none">1. No way of knowing if the sample is typical of the population being studied.2. Can be expensive and time consuming.3. It involves the researcher being dependent on certain individuals.

Table 3.5 - Advantages and disadvantages of snowball sampling.

3.5. REFERENCES

- Brewer, K (2005) The use of volunteers in psychological research Psychology Information for Students 3, 8-11
- Cheurprakobit, S (2000) Police-citizen contact and police performance: Attitudinal differences between Hispanics and non-Hispanics Journal of Criminal Justice 28, 325-336
- Foot, H & Sanford, A (2004) The use and abuse of student participants Psychologist May, 256-259
- Gunnarsson, E (2009) "I think I have had a good life": The everyday lives of older women and men from a lifecourse perspective Ageing and Society 29, 33-48
- Milgram, S (1963) Behavioural study of obedience Journal of Abnormal and Social Psychology 67, 371-378
- Miller, G et al (2007) Ovulatory cycle effects on tip earnings by lap dancers: Economic evidence for human oestrus Evolution and Human Behaviour 28, 6, 375-381
- Nation, J.R (1997) Research Methods Upper Saddle River, NJ: Prentice Hall

4. DESCRIPTIVE STATISTICS: AVERAGES AND MEASURES OF CENTRAL TENDENCY

- 4.1. Introduction
- 4.2. Mean
- 4.3. Median
- 4.4. Mode
- 4.5. Levels of data and measures of central tendency
- 4.6. Measures of central tendency and data from famous research in psychology
- 4.7. References

4.1. INTRODUCTION

Averages are an easy way to summarise the data, but being averages they mask trends and variations in the large set of scores (Whatson 2006)(table 4.1).

The measures of location or central tendency are used to show where in the main a group of scores is located.

More formally, the central tendency is "the value in a group of values which is the most typical for the group or the score which all other scores are evenly clustered around" (Coolican 1990 p139).

There are three measures commonly used (mean, median, and mode), and the correct choice is important or else a false picture of the results may be produced.

ADVANTAGES

1. Summarise data into meaningful patterns for the reader.
2. Individual scores are hard to interpret with vast sets of data.
3. Allows the easy to see comparison between groups.
4. Possible to estimate scores in the population based on measures of central tendency of the sample studied - known as "point estimates" of the population (DSE212 Course Team 2007).

DISADVANTAGES

1. Mask trends and variations in data.
2. Gives no sense of individual scores, particularly in large set of data.
3. Summary data can appear different depending on the measure of central tendency used.
4. Different measures of central tendency are sensitive to different aspects of the data including the level of data (DSE212 Course Team 2007).

Table 4.1 - Advantages and disadvantages of averages and measures of central tendency to summarise data.

4.2. MEAN

This is calculated by adding up all the scores and dividing by the number of scores.

Example 1:

$$(A) \quad 1, 2, 3, 4, 5 = 15/5 = 3$$

$$(B) \quad 15, 22, 10, 16, 45 = 108/5 = 21.6$$

ADVANTAGES

1. Used to show if a particular score is above or below the average for the group, like a fulcrum of a balanced pair of scales (Coolican 1990).

In example 1 (A), two scores are below the mean, two above, and one score is average. But in (B) three scores are below the mean, and two are above in varying amounts.

2. Good measure to compare performance of two groups on the same test.

In example 1, group (B) scores much higher than group (A), if these are test scores. An inferential statistics test would tell us if the difference is significant ¹.

3. The only measure of central tendency for parametric inferential statistics tests ².

4. It is the most sensitive measure of central tendency.

5. It takes account of the total and all the individual scores in its calculation. Thus more sophisticated than the mode.

6. Best with interval and ratio data.

7. Useful with continuous data (ie: not just whole numbers).

8. Used with normal or even distribution of data.

DISADVANTAGES

1. Not applicable with nominal and ordinal data.

2. Because it is a sensitive measure, it can change violently; eg: combining scores (A) and (B) in example 1 gives a mean of 12.3.

3. Not good if data set includes small numbers of extreme scores (outliers or outriders).

Example 2:

$$(A) \quad 5, 6, 8, 8, 9, 12, 92 = 20$$

¹ Significant = results could not have occurred by chance alone.

² Parametric tests are a more powerful type of inferential statistics test.

(B) median = 8

(C) remove outlier (92) and mean = 8

4. Not good if scores vary greatly as in example 1 (B).
5. Cannot be calculated without knowing all the individual scores.
6. Does not necessarily show middle point of data (ie: when 50% above and 50% below).
7. Can produce a meaningless figure in certain situations, like when whole numbers required.

Example 3:

(A) Number of children in 5 families: 1, 1, 2, 3, 4 = mean 2.2 children per family.

(B) median = 2; mode = 1 (more appropriate)

8. The mean calculated is not necessarily the same as an actual score (Moxon et al 2003).

Example 4:

IQ test scores: 100, 101, 102, 109, 116 = 105.6 (not actual IQ score as only whole numbers used)

Table 4.2 - Advantages and disadvantages of the mean as a measure of central tendency.

4.3. MEDIAN

This is calculated by putting the data into numerical order (ranking) and taking middle one (if odd number of scores) or halfway between middle ones (if even number of scores).

Example 5:

(A) 4, 12, 13, 19, 21 = 13

(B) 4, 12, 13, 19, 21, 26 = between 13 and 19 = 16

ADVANTAGES

1. Shows exact middle point of the data with 50% of scores above and 50% below.
2. Easier to calculate than the mean (Coolican 1990).
3. Unaffected by extreme values in one direction (ie: outliers) (Coolican 1990).

Example 6:

- (A) 5, 6, 8, 8, 9, 12, 92 = 8
- (B) 5, 6, 8, 8, 9, 12 = 8

4. Best measure of central tendency for ordinal data, but can also be used with interval and ratio data.
5. Best for highly skewed distribution of data.
6. Can be calculated where extreme values unknown (Coolican 1990).
7. Best with a few high and a few low scores.
8. More sophisticated measure than the mode.

DISADVANTAGES

1. Can be unrepresentative if few widely distributed values.

Example 7:

- (A) 1, 2, 3, 98, 99 = 3
- (B) mean = 40.1 (better)

2. Does not take exact value of each score as in the mean (Coolican 1990).
3. Not suitable with parametric inferential statistics tests.
4. Can be affected by slight change in data (Cox 2000).

Example 8:

- (A) 1, 2, 68, 69 = 35
- (B) 1, 2, 3, 68, 69 = 3 (now)

5. Time consuming to rank large sets of data.
6. Gives less information about the data than the mean.
7. Not as good as the mean for comparing two groups on the same test.

Example 9:

Using data from example 1, median for (A) = 3 and for (B) = 16

8. Where there are an even number of scores, the median may not be an actual score.

Example 10:

5, 6, 12, 15 = between 6 and 12 = 9

Table 4.3 - Advantages and disadvantages of the median as a measure of central tendency.

4.4. MODE

This is the most common score in the data set.

Example 11:

- (A) 1, 2, 3, 3, 4, 5 = 3 (unimodal - single mode)
- (B) 1, 1, 2, 3, 3, 4, 5 = 1 and 3 (bimodal - two modes)
- (C) 1, 1, 2, 2, 3, 3, 4, 5 = 1, 2 and 3 (multimodal)

ADVANTAGES

1. Shows score obtained by more members of the group; ie: most frequent occurrence.
2. Unaffected by extreme values (outliers) in one direction (Coolican 1990).
3. Can be calculated when extreme values unknown (Coolican 1990).
4. Easy to calculate.
5. Good for skewed distributions.
6. Only measure of central tendency for bimodal and multimodal distributions.
7. Best for nominal data, but can be used with ordinal, interval and ratio data.
8. Mode is always an actual score from the data.

DISADVANTAGES

1. Does not take exact value of each score as in the mean (Coolican 1990).
2. Cannot be used with parametric inferential statistics tests.
3. Not useful with small sets of data with many modes (Coolican 1990).

Example 12:

1, 2, 3, 4 = no mode

4. Crude and unsophisticated measure of central tendency.
5. Can change violently with added data.

Example 13:

- (A) 1, 2, 2, 3 = 2 (median = 2)
- (B) 1, 2, 2, 3, 14, 14, 14, = 14 (median = 3; mean = 7.1)

6. Not useful for dichotomous nominal variables (ie: two categories only).
7. Does not necessarily show mid-point of data.

8. Not useful for data grouped together as grouping can change the mode (Coolican 1990).

Example 14:

(A) Ages (yrs): 27, 31, 35, 38, 39, 41, 45

(B) Grouped as: 25-30 - 1; 31-35 - 2; 36-40 - 3; 41-45 - 1;
mode = 36-40 category

(C) Grouped as: 25-39 - 5; 40+ - 2; mode = 25-39

Table 4.4 - Advantages and disadvantages of the mode as a measure of central tendency.

4.5. LEVELS OF DATA AND MEASURES OF CENTRAL TENDENCY

The four levels or types of data are:

- Nominal - numbers used to distinguish categories;
- Ordinal - same as nominal data but in order;
- Interval - a scale with equal distance between the units, but no fixed zero point;
- Ratio - a scale with equal distance between the units and fixed zero point.

Not all measures of central tendency can be used with all levels of data (table 4.5).

Participant number	Age (yrs)	Age category	Likert scale attitude	Food choice
1	27	3	2	1
2	41	3	2	1
3	35	3	1	2
4	28	3	4	2
Nominal data	Interval/ ratio data	Ordinal data 1 = 0-16 2 = 17-25 3 = 26+ yrs	Ordinal data 1 = strongly disagree 2 = disagree 3 = unsure 4 = agree 5 = strongly	Nominal data 1 = chocolate 2 = crisps 3 = neither
MODE none	none	3	2	2
MEDIAN 3	28	3	2	2
MEAN	29.2	2.6	2.8	1.8

(**Bold** = most useful summary of data; *italics* = least useful or meaningless)

Table 4.5 - Examples of different levels of data and which measure of central tendency best.

4.6. MEASURES OF CENTRAL TENDENCY AND DATA FROM FAMOUS RESEARCH IN PSYCHOLOGY.

1. Mode not useful

Eysenck (1952) compared the studies of patients improving or being cured by psychoanalysis and eclectic therapies to see which was more effective.

The mode is not useful to summarise the data, while the median and mean are better (table 4.6).

	Psychoanalysis (5 studies)	Eclectic therapy (19 studies)
% cured, much improved, or improved	39, 62, 47, 50, 67	46, 41, 69, 64, 50, 55, 68, 77, 58, 61, 54, 53, 63, 66, 51, 73, 70, 50, 68
Mode Median	none 50	50 and 58 58

(Data from Eysenck 1952 in Gross 1990 p341)

Table 4.6 - Data from Eysenck (1952).

2. Mean not useful

Stanley Milgram (1963), in his experiment on obedience, recorded the maximum voltage that participants stopped obeying at. If the measures of central tendency are calculated, the mean is not useful here because it suggests a figure of voltage that nobody stopped at. The median and mode are more accurate because the data are positively skewed (ie: mode and median greater than the mean)(Coolican 1990)(table 4.7).

300 volts	5
315	4
330	2
345	1
360	1
375	1
Mode Median	450 450

(Data from Milgram 1963 in Gross 1990 p123)

Table 4.7 - Number of participants who stopped obeying at each voltage.

3. Mean is best

McGinnies (1949) compared an individual's speed of recognising neutral and taboo words presented very quickly. The mean is the best measure of central tendency because interval/ratio data have been collected. The median and the mode are less useful (table 4.8).

Participant Number	Neutral Words	Taboo Words
1	55	184
2	44	94
3	54	80
4	103	126
5	40	64
6	70	130
7	57	104
8	63	76
9	59	130
10	49	223
11	46	77
12	57	91
13	33	37
14	34	54
15	46	56
Mode	46 and 57	130
Median	51.5	85.5

(Data adapted from McGinnies 1949 in Gross 1990 p14)

Table 4.8 - Time to recognise neutral and taboo words (milliseconds).

4. None of the measures of central tendency useful

Dement and Kleitman (1957) investigated the recall of dreams or not by nine participants woken during REM and NREM sleep. None of the measures of central tendency are that useful here. The median and mean produce part numbers when only whole numbers are used (ie: recalled dream or not)(table 4.9). This is nominal data and the mode is usually best. The authors simply calculated the totals for each column.

5. Different measures produce different results

In Asch's (1956) experiment on conformity to the majority, he recorded the number of times (out of 12 trials) participants conformed to the group majority's wrong answer (experimental group) compared to doing the task alone (control group).

Participant number	REM sleep: dream recalled	REM sleep: not recalled	NREM sleep: dream recalled	NREM sleep: not recalled
1	17	9	3	21
2	26	8	2	29
3	36	4	3	31
4	37	5	1	34
5	24	6	2	23
6	4	1	0	5
7	2	2	0	2
8	2	1	0	1
Mode	2 and 4	1	0	none
Median	10.5	3.5	0.5	13

(Data from Dement and Kleitman 1957 in Gross 1990 p242)

Table 4.9 - Data from Dement and Kleitman (1957).

The mode is not useful because it is the same for both groups. The median gives a whole number, which is more appropriate, as individuals can only conform or not conform. The mean is a part number (ie: includes a fraction)(table 4.10). Despite Asch's study being remembered for the conformity, the average showed that conformity was quite low out of 12 trials.

Number of times conformed	Control group (n = 37)	Experimental group (n = 123)
0	35	29
1	1	8
2	1	10
3		17
4		6
5		7
6		7
7		4
8		13
9		6
10		6
11		4
Mode	0	0
Median	0	3

(Data from Asch 1956 in Friend et al 1990)

Table 4.10 - Distribution of participants based on number of times conformed to wrong answers.

4.7. REFERENCES

Asch, S.E (1956) Studies of independence and conformity: A minority of one against a unanimous majority Psychological Monographs 70, 9, whole issue

Coolican, H (1990) Research Methods and Statistics in Psychology London: Hodder & Stoughton

Cox, E (2000) Psychology for AS Level Oxford: Oxford University Press

Dement, W & Kleitman, N (1957) The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming Journal of Experimental Psychology 53, 5, 339-346

DSE212 Course Team (2007) Exploring Psychological Research Methods Milton Keynes: Open University

Eysenck, H.J (1952) The effects of psychotherapy: An evaluation Journal of Counselling Psychology 16, 319-324

Friend, R et al (1990) A puzzling misinterpretation of the Asch "conformity study" European Journal of Social Psychology 20, 29-44

Gross, R.D (1990) Key Studies in Psychology London: Hodder & Stoughton

McGinnies, E (1949) Emotionality and perceptual defence Psychological Review 56, 244-251

Milgram, S (1963) Behavioural study of obedience Journal of Abnormal and Social Psychology 67, 371-378

Moxon, D; Brewer, K & Emmerson, P (2003) Heinemann Psychology AS for AQA A Oxford: Heinemann

Whatson, T (2006) Beyond Reasonable Doubt: A Guide to Designing Experiments in the Behavioural Sciences Milton Keynes: Open University

5. CROSS-SECTIONAL LONGITUDINAL STUDIES: THE BEST OF BOTH WORLDS

- 5.1. Introduction
- 5.2. Kleinspehn-Ammerlahn et al (2008)
- 5.3. References

5.1. INTRODUCTION

Traditionally there is a distinction between cross-sectional studies, which compare different groups of people at one point in time (eg: men vs women), and longitudinal studies that test the same group of participants at different points in time (eg: at age 20 and age 40).

The cross-sectional longitudinal study combines these methods and studies different groups at different points in time. An example of this method is Kleinspehn-Ammerlahn et al (2008).

5.2. KLEINSPEHN-AMMERLAHN ET AL (2008)

This study was interested in older adults' self-perception of ageing using participants from the Berlin Ageing Study (BASE)(Baltes and Mayer 1999) ³.

Data were collected at four points in time (the longitudinal element of the study):

- Time 1 (T1) - between 1990-3
- T2 - 1993-4
- T3 - 1995-6
- T4 - 1997-8

The number of participants was reduced at each time point, mainly due to death. For example, T1 had 516 participants, and T4 only 132 of them.

The cross-sectional element of the study involved the division of the participants into gender groups and six age groups (70-74, 75-79, 80-84, 85-89, 90-94, 95+ at T1)(figure 5.1). Originally there were 43 men and 43 women allocated to each age group.

³ <http://www.base-berlin.mpg.de>

CROSS-SECTIONAL LONGITUDINAL STUDY

Longitudinal element:	T1 1990-3	T2 1993-4	T3 1995-6	T4 1997-8	
Cross-sectional element:					
1. 70-74 male	→	→	→	→	76-80
2. 70-74 female	→	→	→	→	76-80
3. 75-79 male	→	→	→	→	81-85
4. 75-79 female	→	→	→	→	81-85
5. 80-84 male	→	→	→	→	86-90
6. 80-84 female	→	→	→	→	86-90
7. 85-89 male	→	→	→	→	91-95
8. 85-89 female	→	→	→	→	91-95
9. 90-94 male	→	→	→	→	96-100
10. 90-94 female	→	→	→	→	96-100
11. 95+ male	→	→	→	→	101+
12. 95+ female	→	→	→	→	101+

LONGITUDINAL STUDY

One age group at different points - eg: 70-74 male at T1, T2, T3, T4

CROSS-SECTIONAL STUDY

Comparison of all 12 groups at T1 or T2 or T3 or T4

Figure 5.1 - Representation of cross-sectional longitudinal study as used by Kleinspehn et al (2008) compared to the longitudinal and the cross-sectional study.

The researchers measured the following variables:

a) Subjective or felt age - using two items: "How old do you feel?", and "How old do you feel when you look at yourself in a mirror?". Scoring was based on the age given minus the chronological age. This produced a discrepancy score. A minus number meant that the individual felt younger than their actual age, and a positive number that they felt older.

b) Satisfaction with ageing - using the Attitudes Toward Own Ageing subscale from the Philadelphia Geriatric Center Morale Scale (Lawton 1975) which contains five items, like "Things keep getting worse as I get older" and "I am as happy now as I was when I was

younger", measured on a five-point scale (from 1 = low agreement to 5 = high agreement).

c) Individual differences - including socio-economic status (SES)(based on income, occupation, and education), number of illnesses, cognitive functioning (on fourteen standardised tasks), and loneliness. The latter was measured with the five-point scale of four items of UCLA Loneliness Scale (Russell et al 1980)(eg: "There are people I feel close to", and "I feel part of a group of friends").

The results showed that participants generally felt younger than their actual age (by average of 12.86 years), and showed high levels of satisfaction. Over the period of the study (approximately six years) there was a small decline in these measures as shown by the mean (table 5.1).

Measures:	T1	T4
Felt age discrepancy (yrs) ("How old do you feel?")*	-12.96	-12.54
Physical age discrepancy (yrs) ("How old do you feel when you look at yourself in the mirror?")*	-9.51	-6.84
Satisfaction with ageing (T score)**	50.00	45.34

(* = minus number means feel younger than chronological age)

(** = T scores are standardised scores based on T1)

(After Kleinspehn-Ammarlahn et al 2008)

Table 5.1 - Mean scores for all participants over the course of the study (longitudinal element).

In terms of the cross-section, males were more satisfied with ageing than women. Participants in the older age groups showed a greater decline in ageing satisfaction but an increase in the discrepancy of felt age (ie: feeling younger) compared to younger age groups. Greater satisfaction with ageing was linked to less illnesses, lower social loneliness, higher cognitive functioning, and higher SES than the average.

The cross-sectional longitudinal study method has advantages and disadvantages common and different to the separate use of the longitudinal study and the cross-sectional study (table 5.2).

ADVANTAGES

1. Advantages in common with cross-sectional studies:
 - i. Compares groups.
 - ii. Provides average for different groups.
2. Advantages in common with longitudinal studies:
 - i. Overcome cohort effect. This is where groups can be different due to differences in the group themselves (eg: life experiences growing up of seventy vs ninety year-old). In this study all German participants (ie: common historical experience).
 - ii. Compares individual scores.
 - iii. Can observe change over time.
 - iv. Shows sequence of events.
3. Like an experiment combining independent and related designs.
4. Better than cross-sectional study because control for cohort effects, and follows individuals as well as having group averages.
5. Better than longitudinal study because large number of participants overcomes risk of high drop-out with a small sample.
6. Random sampling of participants from 1908 older persons in Berlin city registry.
7. The interviews took place at the participants' place of residence, face-to-face, with trained research assistants who read out the questions.
8. There was a good time period between each set of questions (ie: two years approximately), so the participants would not become bored with being studied too regularly.
9. The researchers were sensitive to the older individuals' needs; eg: severely visually impaired not asked how old they looked in the mirror. This partly accounted for differences in the number of respondents on certain measures at each time period.
10. Questionnaires with established reliability and validity were used; eg: UCLA Loneliness Scale Cronbach's alpha (measure of reliability) = 0.70.

DISADVANTAGES

1. Disadvantages in common with cross-sectional studies:
 - i. Problem of matching participants in each group.
 - ii. Larger number of participants required than longitudinal study.
2. Disadvantages in common with longitudinal studies:
 - i. Problem of participant drop-out: T1 = 516 participants, T2 = 361, T3 = 206, and T4 = 132.
 - ii. Time consuming (ie: six-eight years).

iii. Participants may be affected by repeated testing: each set of testing took a number of sessions, and sessions lasted 90 minutes.

iv. Any methodological mistakes at beginning of study hard to rectify later.

3. Difficult to replicate because of vast amount of data collected, cost, and complexity of design.

4. It was only a "short-term longitudinal study".

5. Study based in one city, and participants may or may not have lived there all their lives.

6. Because of the loss of participants by T4, the sample is not as representative as at T1 (eg: more younger participants). The oldest-old (ie: over 85s) are "survivors" who are often in better health than the average and more satisfied with their lives (Perls 1995).

7. Those who agreed to take part are self-selecting, and may not be typical of the 1392 people who did not.

8. The authors admitted that more measurement points and longer time periods were needed.

9. Not all individuals participated on all four measurement occasions, and not all measures were taken at each occasion (eg: subjective age not measured at T2). This meant that there were different numbers of participants for different measures. For example, at T3, felt age discrepancy, n = 205 and physical age discrepancy, n = 200. At T4 it was 132 and 129 respectively.

10. The use of self-reported questionnaires does depend upon the honesty/accuracy of replies either through direct lying, memory problems, or misunderstanding of the questions. Because the study focused upon subjective perception of ageing, objective measures were not taken where applicable (eg: number of friends).

Table 5.2 - Advantages and disadvantages of the cross-sectional longitudinal study method as used by Kleinspehn et al (2008) compared to longitudinal studies and cross-sectional studies.

5.3. REFERENCES

Baltes, P.B & Mayer, K.U (1999)(eds) The Berlin Ageing Study: Ageing from 70 to 100 New York: Cambridge University Press

Kleinspehn-Ammlahr, A; Kotter-Grühn, D & Smith, J (2008) Self-perceptions of ageing: Do subjective age and satisfaction with ageing change during old age? Journal of Gerontology: Psychological Sciences 63B, 6, P377-P385

Lawton, M.P (1975) The Philadelphia Geriatric Center Morale Scale: A revision Gerontology 30, 85-89

Perls, T.T (1995) The oldest old Scientific American January, 50-55

Russell, D et al (1980) The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence Journal of Personality and Social Psychology 39, 472-480

6. STUDYING BEHAVIOUR WITH AN INTERNET-BASED QUESTIONNAIRE: AN EXAMPLE FROM RESEARCH ON SLEEP DISORDERS

- 6.1. Internet-based questionnaires
- 6.2. References

6.1. INTERNET-BASED QUESTIONNAIRES

Girard and Cheyne (2006) set up an online study with individuals who had suffered an episode of "sleep paralysis"⁴ in the previous twelve months (n = 410).

The researchers were able to draw from 1876 individuals who had visited their website, but the majority (89%) were from North America. The sample of respondents was also biased in Caucasian descent (84%), female (71%), age (mean 31.3 years old), and students (23%).

Cheyne (2005) used the Waterloo Unusual Sleep Experiences Scale website (<http://watarts.uwaterloo.ca/~acheyne/spquest01.html>) to investigate the types of hallucinations during sleep paralysis. Between March 2001 and February 2004, there were 12 942 respondents (of which 5799 were used in the study).

These two studies are typical of an increasing number in that the internet is used as the means to collect questionnaire data. There are a number of advantages and disadvantages of using this method (table 6.1).

ADVANTAGES

1. Cheaper than postal or face-to-face questionnaires.
2. Allows anonymity of respondent apart from email address.
3. Wider geographically sample (anywhere in the world).
4. Able to collect data from a very large sample.
5. Individuals can fill out the questionnaire in their own time.
6. Can answer questions which require respondents to check

⁴ Sleep paralysis" is a state of "involuntary immobility" due to the intrusion of wakefulness into REM sleep. The upshot is that individuals perceive themselves as awake but paralysed, when in fact they are dreaming. It is a highly vivid dream, or put another way, it is the "maintenance of REM consciousness when waking" (Girard and Cheyne 2006).

information (eg: in diaries) as there is no time limit as with telephone or face-to-face interviews.

7. A longer questionnaire can be constructed as the individual does not have to complete it in one sitting.

8. Questionnaire can be available for a long time period on a website, and only needs the researcher to check for replies intermittently.

9. Removes interviewer-interviewee interaction variables that are present in face-to-face interviews.

DISADVANTAGES

1. Self-selecting sample of those who visited website and signed up for study.

2. Under-representation of individuals from certain backgrounds; eg: without Internet access, poorer, older.

3. No independent confirmation of information.

4. If questions unclear no interviewer to clarify the meaning.

5. In face-to-face interviews, interviewer may be able to pick up extra information from how respondents behave (eg: hesitation before answering question).

6. As with all questionnaires depend on the honesty, and accuracy of recall of respondents.

7. Internet-based questionnaires better for short and forced-choice answers rather than the longer, qualitative ones.

8. Researchers have no control over situation when and where questionnaire filled in (eg: group of friends together; respondent not concentrating).

9. Some individuals may respond more than once, particularly using a pseudonym. This is a problem with individuals who want to disrupt the research project (eg: the "screw you" effect; Masling 1966).

Table 6.1 - Advantages and disadvantages of Internet-based questionnaires.

Table 6.2 compares Internet-based questionnaires with postal, telephone, and face-to-face interviews, while table 6.3 gives the general problems of questionnaire-based research.

TYPE OF QUESTIONNAIRE	ADVANTAGE OF	DISADVANTAGE OF
Postal	Cheaper and larger	Self-selecting sample
Telephone	Cheaper and larger	Self-selecting sample
Face-to-face	No interviewer-interviewee	Unable to clarify questions and

Table 6.2 - Advantage and disadvantage of Internet-based questionnaires compared to other questionnaire methods.

- No independent confirmation of information
- Honesty of respondents (ie: lying deliberately)
- Recall accuracy of respondents (ie; want to tell truth but cannot remember)
- Socially desirable answers (ie: embarrassed to give true answer) - less with anonymous Internet-based questionnaires
- Misunderstanding of questions
- Type of data usually limited to quantitative

Table 6.3 - General problems of questionnaire research.

6.2. REFERENCES

Cheyne, J.A (2005) Sleep paralysis episode frequency and number, types and structure of associated hallucinations Journal of Sleep Research 14, 319-324

Girard, T.A & Cheyne, J.A (2006) Timing of spontaneous sleep-paralysis episodes Journal of Sleep Research 15, 222-229

Masling, J (1966) Role-related behaviour of the subject and psychologist and its effect upon psychological data. In Levine, D (ed) Nebraska Symposium on Motivation Lincoln, Neb: University of Nebraska Press

7. STUDYING SPECIFIC GROUPS: THE NUN STUDY

- 7.1. Advantages and disadvantages of studying nuns
- 7.2. The Nun Study
- 7.3. References

7.1. ADVANTAGES AND DISADVANTAGES OF STUDYING NUNS

Researchers study specific groups like nuns because they have advantages over samples from the general population.

1. Nuns are an ideal control group with homogeneous lifestyles that eliminate confounding factors. They are usually non-smokers and do not use alcohol/recreational drugs, and have the same marital status, reproductive and family history as well as living and working together (Dowling 2008). The variables being studied can be isolated so much easier with the removal of these factors found in the general population. "Outside a laboratory, it would be hard to find as pure an environment for research" (Snowden 2001).

For example, Butler and Snowden (1996) compared mortality rates between Catholic sisters and the general population between 1965 and 1989. The nuns had much lower rates of smoking-related disease mortality, but higher rates of cancers of the breast and reproductive organs than women in the general population.

2. Being a settled group, the nuns can be studied for long periods without problems of participants moving away or being unable to contact. Participant loss is a major problem in longitudinal studies.

3. They are generally willing to participate in research because of altruistic motives (Dowling 2008). This includes allowing research on the body after death. Using volunteers overcomes any ethical problems related to forced participation in research.

4. Convents tend to have detailed record-keeping which gives extensive membership lists and historical records for researchers (Snowden 2001).

5. It is rare in the general population to find a group of women who have not had children.

There are disadvantages in studying nuns,

particularly in relation to their difference to the general population (table 7.1).

- Smoking, alcohol and recreational drug use
- Diet (eg: less junk food)
- Celibacy/reproductive history
- Marital status/family history
- Contact with men
- Work history
- Nature of convent environment (eg: no television)
- Religious beliefs and behaviour
- Personality characteristics (eg: choosing to become a nun)

Table 7.1 - Areas where nuns might be untypical compared to the general population.

7.2. THE NUN STUDY

The Catholic nuns of the School Sisters of Notre Dame (SSND) at the Good Counsel Hill convent in Mankato, near St. Paul, Minnesota, USA were first approached in 1986 (Snowden 2001), and this became the basis of The Nun Study with other SSND convents included later.

In total 678 nuns between the age of 75 and 102 years have been studied in relation to ageing and Alzheimer's disease. For example, Butler et al (1996) found that cognitive decline, over a two-year period, using the Mini-Mental State Examination, was greater for lower educated nuns (as based on educational qualifications).

Table 7.2 lists some of the key findings in relation to ageing from the Nun Study.

- College graduates live longer while maintaining abilities for self-care (Snowden et al 1989).
- "Low normal" cognitive functioning as predictor of cognitive impairment (Greiner et al 1996).
- Contradictions to traditional ideas about Alzheimer's disease; eg: "Sister Mary" had high cognitive abilities at 101 years old despite neurofibrillary tangles and senile plaque (classic brain damage in Alzheimer's disease) (Snowden 1997).
- Specific apolipoprotein-E APOE gene and high cognitive functioning (Riley et al 2000).
- Positive emotions in early adulthood and living longer (Danner et al 2001).

Table 7.2 - Five key findings on ageing from the Nun Study.

7.3. REFERENCES

Butler, S.M & Snowden, D.A (1996) Trends in mortality in older women: Findings from the Nun Study Journal of Gerontology: Social Sciences 51B, 4, S201-S208

Butler, S.M et al (1996) Age, education and changes in the Mini-Mental State Examination scores of elderly women: Findings from the Nun Study Journal of the American Geriatrics Society 44, 675-681

Danner, D et al (2001) Positive emotions in early life and longevity: Findings from the Nun Study Journal of Personality and Social Psychology 80, 5, 804-813

Dowling, T (2008) The scientific application of nun testing Guardian 7/10, pII.3

Greiner, P.A et al (1996) The loss of independence inactivities of daily living: The role of low normal cognitive function on elderly nuns American Journal of Public Health 86, 1, 62-66

Riley, K.P et al (2000) Cognitive function and apolipoprotein-E in the very old: Findings from the Nun Study Journal of Gerontology: Social Sciences 55B, 2, S69-S75

Snowden, D.A (1997) Aging and Alzheimer's disease: Lessons from the Nun Study The Gerontologist 37, 2, 150-156

Snowden, D.A (2001) Aging With Grace New York: Bantam

Snowden, D.A et al (1989) Education, survival and independence in elderly Catholic sisters American Journal of Epidemiology 130, 999-1012

8. PSYCHOLOGY DEBATE: THIS HOUSE BELIEVES THAT THE SELF IS BEST STUDIED WITH QUANTITATIVE METHODS

FOR

1. Quantitative methods here would include structured questionnaires and tests that give clear, unambiguous scores for different individuals. These scores can be compared to other individuals or to group norms, and places the person in a social context.
2. Many aspects of the self only make sense in relation to others, like higher and lower self-esteem. Qualitative methods do not allow comparison between individuals very easily.
3. Well designed tests have validity and reliability. The reliability means that individuals tested at different times will get the same scores unless their self has changed. Any difference in score will not be a facet of the test.
4. Having a clear measure of the self allows researchers to see how it changes during an experiment. In the example of the "Mr.Clean/Mr.Dirty" experiment (Morse and Gergen 1970), participants' self-esteem in relation to a job interview increased or decreased depending upon the presence of the ideal candidate ("Mr.Clean") or a scruffy candidate ("Mr.Dirty").
5. Accurate measures of the self are useful in counselling and therapy to show improvements and changes.

AGAINST

1. Forced-choice answers, like yes or no, and numerical scores cannot do justice to the meaning and experience of the self.
2. Answers to forced-choice questions can vary depending on the individual's mood and what is happening at the time.
3. The self is made up of many different components like self-esteem and body image, and these can be explored in more depth by qualitative (open-ended) interviews and questionnaires.

Qualitative methods allow the individual to talk about themselves without any researcher-imposed boundaries, and afterwards the researcher can analysis

what has been said.

4. Simple statements similar to "I like myself", yes or no, are prone to social desirability bias. It is harder in a qualitative interview for individuals to always give what they think is the desired answer.

5. With qualitative methods, like the repertory grid technique (RGT)(Kelly 1955), the individual is free to describe themselves without boundaries placed by the researcher. The RGT can show the researcher the worldview of the participant. It is a participant-led methodology.

BIBLIOGRAPHY

Banister, P et al (1994) Qualitative Methods in Psychology
Buckingham: Open University Press

Kelly, G.A (1955) The Psychology of Personal Constructs, volumes 1 and 2 New York: Norton

Morse, S & Gergen, K.J (1970) Social comparison, self-consistency, and the concept of self Journal of Personality and Social Psychology 16, 1, 148-156

9. TWO CLASSIC 1960S SOCIAL PSYCHOLOGY STUDIES ON THE INFLUENCE OF OTHER PEOPLE: LABORATORY VERSUS NATURALISTIC STUDY

- 9.1. Introduction
- 9.2. Latané and Darley (1968): "Smoke incident" laboratory experiment
- 9.3. Milgram et al (1969): Field experiment on conformity to a crowd
- 9.4. Control groups
- 9.5. References

9.1. INTRODUCTION

Experiments can be performed in the laboratory, as is more common, or in a naturalistic setting (field experiments). When studying social behaviour as in the influence of others on the individual, both types of experiment have been used. The laboratory experiment has greater control over variables, but low ecological validity, while the field experiment tends to have the opposite.

9.2. LATANÉ AND DARLEY (1968): "SMOKE INCIDENT" LABORATORY EXPERIMENT

Latané and Darley (1968) noted that:

Most emergencies are, or at least begin as, ambiguous events. A quarrel in the street may erupt into violence, but it may be simply a family argument. A man staggering about may be suffering a coronary or an onset of diabetes; he may be simply drunk. Smoke pouring from a building may signal a fire; on the other hand, it may be simply steam or air-conditioning vapour. Before a bystander is likely to take action in such ambiguous situations, he must first define the event as an emergency and decide that intervention is the proper course of action (p216).

In the process of defining an event as an emergency the reaction of others is important. Are they behaving normally or do they appear to think that an emergency is happening? Latané and Darley (1968) investigated this issue in a classic laboratory experiment that manipulated the reaction of other people to an unexpected event - smoke filling a room.

Two predictions were made by the researchers ⁵:

a) Individuals are less likely to interpret the ambiguous event as an emergency in a group than alone;

b) Individuals are less likely to interpret the ambiguous event as an emergency when the reaction of others is passive than when alone.

In a social situation, individuals are watching others and are aware that they are being watched. Remaining calm is a desirable trait in Western society, and an individual does not want to over-react to an event that is not an emergency: "If each member of a group is, at the same time, trying to appear calm and also looking around at the other members to gauge their reactions, all members may be led (or misled) by each other to define the situation as less critical than they would if alone" (Latané and Darley 1968 p216).

Latané and Darley recruited, by telephone, 58 male students at Columbia University, New York, to be interviewed about "some problems involved in life at an urban university". This was a volunteer sample based on those living in campus residences.

Many of the classic studies in psychology are based upon male students. Students have been used in around 70% of recent studies in personality and social psychology (Kimmel 1996). It has been argued that student participants do not actually mind being deceived (Kimmel 1998). On the other hand, students may deceive experimenters about their naivety to specific tests or pass information about the study to peers who may or may not be participants as well (Tindell and Bohlander 2005).

Not only are students (undergraduate and postgraduate) not typical of the general population, but the exclusion of female participants promotes the behaviour of men as the norm (Foster 1997). An untypical sample of participants limits the generalisability of the findings to the general population.

The response rate of individuals who agree to participate in the research over the telephone and whether they turn up at the experiment is an important concern. In relation to medical research, Junghans et al (2005) compared the "opt-in" and "opt-out" methods of recruitment of volunteers. In the "opt-in" group, individuals were approached and asked by telephone to

⁵ These predictions are the experimental (or alternate) hypotheses which are being tested by the experimenters. They are predictions about what is expected to happen, and will be supported or not by the data collected.

volunteer for a study, and they had to contact the researchers again to agree to take part. The response rate was 38%. The "opt-out" condition followed the same procedure, but the potential volunteers had to contact the researchers again to ask not to participate in the study. The response rate for participation in the study was 50% here. But this probably only works for studies where the researcher goes to the volunteer. For example, you will be contacted in two weeks if you have not opted-out of the research.

Arriving for the interview, participants were put in a "waiting room" to fill out a questionnaire. During this time, smoke was puffed into the room through a wall vent, and the response of the participant was observed through a one-way glass. "The 'smoke' was finely divided titanium dioxide produced in a stoppered bottle and delivered under slight air pressure through the vent. It formed a moderately fine-textured but clearly visible stream of whitish smoke. For the entire experimental period, the smoke continued to jet into the room in irregular puffs. By the end of the experimental period, vision was obscured by the amount of smoke present" (p217).

The experimental manipulation related to who was in the waiting room with the participant and there were three conditions ⁶:

- Participant alone
- Participant with two confederates ⁷ who were instructed to show no reaction to the smoke ("two passive confederates condition"): "The confederates attempted to avoid conversation as much as possible. Once the smoke had been introduced, they stared at it briefly, made no comment, but simply shrugged their shoulders, returned to the questionnaires and continued to fill them out, occasionally waving away the smoke to do so. If addressed, they attempted to be as uncommunicative as possible and to show apparent indifference to the smoke. 'I dunno,' they said, and no subject persisted in talking" (p217).
- Three naive participants who arrived at the experiment separately, and did not know each other.

⁶ There was no control group; eg: participants in waiting room and no smoke pumped in.

⁷ Confederates (or stooges) are individuals in an experiment who appear as participants but are following instructions from the experimenter. Their presence gives the experimenter the ability to control and manipulate aspects of the experimental situation, but it does mean that the real participants are deceived.

The ideal participants are naive to the purposes of the experiment, and so, it is assumed, they will behave normally. However, participants may try to guess the purpose of the experiment and alter their behaviour to please the experimenter ("demand characteristics"; Orne 1962) or to go against the perceived purpose ("the screw you effect"; Masling 1966). Orne (1962) argued that the participant "must be recognised as an active participant in any experiment, and that it may be fruitful to view the psychology experiment as a very special form of social interaction" (p783).

There were two independent variables of this study could be viewed as two separate experiments. One independent variable was the presence or absence of passive confederates, and the other was the presence or absence of other people.

Six minutes was allowed for the participant to report the presence of smoke to someone outside the room, and then the experiment was terminated. Allowing six minutes for response was the operationalisation of the concept of interpreting the ambiguous event as an emergency. Operationalisation gives an objective scoring system for the experiment. This was also the dependent variable.

In the alone condition, 18 of 24 participants reported the smoke, and the median time was two minutes. With the "two passive confederates condition", only 1 of 10 did, while "other 9 stayed in the waiting room as it filled up with smoke, doggedly working on their questionnaire and waving the fumes away from their faces. They coughed, rubbed their eyes, and opened the window – but they did not report lie smoke" (p218). Obviously there was a significant difference between the two conditions ($p < 0.002$; Fisher's exact test). This finding supported prediction (b) above.

In the "three naive bystanders condition", 24 participants were divided into eight runs of the experiment. Only three participants reported the smoke (38%)(ie: 3 of 8 groups), and this is significantly less than the alone condition ($p < 0.01$; Fisher's exact test). Prediction (a) above was supported.

Not only did more participants lone report the smoke, but they did so quicker than the groups of three. For example, 75% of participants alone had responded by four minutes and only one person from the "three naive participants condition". The median speed of noticing the smoke as viewed through the one-way glass also varied – five seconds when alone and twenty seconds in the group ($p < 0.01$; chi-square).

Table 9.1 summarises the results.

CONDITION	NUMBER OF	% (n) RESPONDING
Alone	24	75 (18)
Two passive	10	10 (1)
Three naive	24 in 8 groups	38 (3)

Table 9.1 - Percentage and number of participants responding within six minutes of smoke appearing in the waiting room.

After six minutes, the participants who had not responded were invited into the interview by an interviewer who made no reference to the smoke, and simply asked if the participant "had experienced any difficulty while filling out the questionnaire". "By this point most subjects mentioned the smoke. The interviewer expressed mild surprise and asked the subject to tell him what had happened. Thus each subject gave an account of what had gone through his mind during the smoke infusion" (p219).

Among the participants who had not responded to the smoke:

..they hit upon an astonishing variety of alternative explanations, all sharing the common characteristic of interpreting the smoke as a non-dangerous event. Many thought the smoke was either steam or air-conditioning vapours, several thought it was smog, purposely introduced to simulate an urban environment, and two (from different groups) actually suggested that the smoke was a "truth gas" filtered into the room to induce them to answer the questionnaire accurately. (Surprisingly, they were not disturbed by this conviction.) Predictably, some decided that "it must be some sort of experiment" and stoically endured the discomfort of the room rather than overreact (p219).

Latané and Darley concluded that this research showed that the "failure to intervene may be better understood by knowing the relationship among bystanders rather than that between a bystander and the victim" (p221).

There are ethical concerns about this experiment:

- Deception - about purpose of experiment; confederates appearing to be other participants; post-smoke interview; covert observation through one-way glass.
- Stress to participants of smoke filling the room.

- Embarrassment at realising part of experiment.
- A debriefing was not mentioned, but it is assumed that the participants were told.
- No informed consent for actual experiment that took place. Were the participants given the opportunity to withdraw their data retrospectively?
- Right to withdraw limited.

Table 9.2 gives the key strengths and weaknesses of this laboratory experiment.

STRENGTHS	WEAKNESSES
1. Control over the variables and participants. 2. Ability to manipulate the research environment. 3. isolation of confounding variables like influence of people not involved in	1. Artificial situation, in many respects, that is not typical of everyday life. 2. Use of deception. 3. Risk of participants altering their behaviour because they guessed it was an experimental

Table 9.2 - Key strengths and weaknesses of the laboratory experiment as used by Latané and Darley (1968).

9.3. MILGRAM ET AL (1969): FIELD EXPERIMENT ON CONFORMITY TO A CROWD

The influence of the behaviour of others can be seen in a naturalistic study of crowd behaviour by Milgram et al (1969). The researchers were interested to see how passersby on a busy city street in New York would respond to a small crowd stood on the pavement and looking up at the window of a nearby building. Would the passersby conform to the behaviour of the crowd looking up?

On two winter afternoons, a 50-foot length of pavement was observed in thirty one-minute trials where confederates would form groups of different sizes ("stimulus crowd") and gaze for 60 seconds at a 6th floor window. The stimulus crowd varied between 1, 2, 3, 5, 10, and 15 persons. During the trial, filming was made to see how many pedestrians passed-by, and how many looked up and/or stopped. The total sample was counted as 1424 pedestrians from the film.

The larger the stimulus crowd, the more passersby looked up and stopped ($p < 0.001$; ANOVA). For example, when the stimulus crowd was one person, 4% of passersby

stopped, while 40% stopped with a stimulus crowd of fifteen. One person induced 42% of passersby to look up, while 15 persons induced 86% ($p < 0.001$; ANOVA).

Overall, "There is some logical basis for joining larger crowds: all other things being equal, the larger the crowd the more likely its members are attending to a matter of interest" (Milgram et al 1969 pp81-82).

This relatively simple study was a field experiment where the independent variable was the size of the stimulus crowd, and the dependent variable was passersby looking up or stopping. There are a number of strengths and weaknesses to this study compared to a laboratory experiment (table 9.3).

STRENGTHS

1. Takes place in a real-life situation (high ecological validity).
2. Not possible or very difficult to perform this study in a laboratory/controlled environment.
3. Researchers still able to manipulate independent variable.
4. Participants assumed to be naive to fact that it was an experiment.
5. Filming pedestrians allowed the researchers to review later, and independent judges were used to count passersby who had looked up or stopped.
6. A large number of people sampled (in a short time).
7. Standardised procedures used, particularly in relation to length of gaze by stimulus crowd: "At a signal, flashed from the sixth-floor window of an office building across the street from this area of sidewalk, a group of confederates (stimulus crowd) entered the middle of the observation area, stopped, and looked up at the sixth-floor window. This gaze was maintained for 60 seconds. At the end of this period the group was signalled to disperse. After the area was cleared of the gathered crowd the procedure was repeated using a different size stimulus crowd. Five randomly ordered trials were conducted for each of the six different size stimulus crowds" (p80).

WEAKNESSES

1. Much less control over variables than laboratory experiment; eg: who passersby; temperature; weather; other distractions in the street.
2. Use of confederates as the stimulus crowd meant that the passersby were deceived.
3. There was no way of gaining informed consent beforehand, no debriefing afterwards, nor right to non-participant and withdraw.
4. Conformity was operationalised as looking up in same direction as stimulus crowd and stopping. This definition has problems like the

length of stopping, and how to confirm exact direction of gaze. Is it a valid definition?

5. A self-selecting sample based on those who at the place of the study at that time. Not possible to know if it was typical of the city or the general population.

6. The use of film only captures the individuals within the camera's range.

7. No control group - ie: how many passersby looked up or stopped without any stimulus crowd.

Table 9.3 - Strengths and weaknesses of the field experiment as used by Milgram et al (1969).

9.4. CONTROL GROUPS

Both experiments described here did not have a control group. Often the control group in psychology experiments or patient research is a "mere afterthought": "Journal articles often go into great detail about the lengths the authors went to in order to recruit their patient groups, but control participants are usually given a brief mention as having been recruited 'through advertising' or 'a volunteer pool'" (Wicks 2007 p22).

When funding for research is limited, spending money on a control sample can be seen as a luxury, particularly if the control condition appears unnecessary. Or it could mean that a control group is made up of those easily available and is not matched for age or sex, for example, with the experimental group(s) (Wicks 2007).

A comparable control group aids the internal validity of the experiment. Internal validity is where the findings "can be attributed with little or no ambiguity to effects of the independent variable" (Kazdin 2003). In other words, it is a confounding variable that has been controlled by the experimenter.

Table 9.4 outlines some examples of control groups and table 9.5 gives the advantages and disadvantages of having a control group in an experiment.

STUDY	INDEPENDENT VARIABLE	CONTROL GROUP
Noise and recall	Level of noise	Memory test in silence
Imitating aggression from TV programme	Level of aggression on television programme	Watch television programme with no aggression, or do not watch any programme, and then level of aggression
Latané and Darley (1968)	Presence/reaction of others to smoke	Fill out questionnaire without smoke either
Milgram et al (1969)	Size of stimulus	No stimulus crowd

Table 9.4 - Examples of control groups in different experiments.

ADVANTAGES	DISADVANTAGES
1. Comparison group for experimental group(s). 2. Gives baseline measures of behaviour before independent variable manipulated in repeated measures experimental design. 3. Able to highlight the effect of the independent variable more strongly when the dependent variable is small or non-existent in the control group.	1. Participants may be asked to perform dull tasks when no manipulation of the independent variable. There is also the risk of order effects in a repeated measures experimental design. 2. Difficulty and cost of recruiting participants for conditions where little is happening. 3. Participants who know they are in the control group may be demotivated or change their

Table 9.5 - Advantages and disadvantages of having a control group in an experiment.

9.5. REFERENCES

Foster, A (1997) Scientific research which excludes women Psychology Teaching October, 37-45

Junghans, C et al (2005) Recruiting patients to medical research: Double blind randomised trial of "opt-in" versus "opt-out" strategies British Medical Journal 331, p940

Kazdin, A.E (2003) Research Design in Clinical Psychology (4th ed) Needham Heights, MA: Allyn & Bacon

Kimmel, A.J (1996) Ethical Issues in Behavioural Research Cambridge, MA: Blackwell

Kimmel, A.J (1998) In defence of deception American Psychologist 53, 803-804

Latané, B & Darley, J (1968) Group inhibition of bystander intervention in emergencies Journal of Personality and Social Psychology 10, 3, 215-221

Masling, J (1966) Role-related behaviour of the subject and psychologist and its effect upon psychological data Nebraska Symposium on Motivation 14, 67-103

Milgram, S; Bickman, L & Berkowitz, L (1969) Note on the drawing powere of crowds of different size Journal of Personality and Social Psychology 13, 2, 79-82

Orne, M.T (1962) On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications American Psychologist 17, 776-783

Tindell, D & Bohlander, R (2005) Participants' naivety and confidentiality in psychological research Psychological Reports 96, 963-969

Wicks, P (2007) Linchpin or afterthought? Psychologist 20, 1, 22-25

10. CARVING UP CONTINUOUS DATA

- 10.1. Introduction
- 10.2. Carving up data
- 10.3. References

10.1. INTRODUCTION

Quantitative data come in two forms - continuous (numbers along a scale, like time) or discrete (separate categories)(table 10.1).

Continuous measurements can take any value on the scale from the highest to the lowest, while discrete measurements take up only certain values. For example, in a memory test, "time taken to recall a word list in seconds" is continuous data because any number is theoretically possible (eg: 10.3secs; 6.5secs). But "number of words recalled correctly" is discrete data because certain numbers are possible (eg: 12 words; 3 words)(DSE212 Course Team 2007).

Nominal and ordinal scales are discrete, whereas interval and ratio data can be continuous or discrete.

CONTINUOUS

How old are you? __ years

DISCRETE

How old are you? 16-25
 26-35
 36-45
 46-55
 55+

Table 10.1 - Example of continuous and discrete measurements.

10.2. CARVING UP DATA

Continuous data can be converted into discrete data, but not the other way around. This is known as "carving" the data, and "the results of an investigation may vary, depending on what sort of cutpoints are chosen for the data carving" (Owen and Froman 2005 p498). Table 10.2 gives an example of the reduction of body mass index (BMI) to a "crude ordinal scale" by researchers. "Such data reduction violates the obvious: People are not merely big or not big" (Owen and Froman 2005 p496).

BMI = weight (kg)/height (m ²)	Continuous data
REDUCED TO*:	
under 18.5 (underweight)	Discrete data
18.5 - 24.9 (normal)	
25.0 - 29.9 (overweight)	
30.0+ (obese)	
REDUCED FURTHER**:	
<25.0 (not overweight)	Discrete data
>25.0 (overweight)	

* Center for Disease Control and Prevention (2005) BMI calculator; quoted in Owen and Froman (2005)

** Lee and Paffenbarger (2000)

Table 10.2 - Example of data carving with BMI (Owen and Froman 2005).

10.3. REFERENCES

DSE212 Course Team (2007) Exploring Psychological Research Methods
Milton Keynes: Open University

Lee, I.M & Paffenbarger, R.S (2000) Associations of light, moderate and vigorous intensity physical activity with longevity American Journal of Epidemiology 151, 293-299

Owen, S.V & Froman, R.D (2005) Why carve up your continuous data? Research in Nursing and Health 28, 496-503